DOCUMENT RESUME

ED 457 198 TM 033 296

AUTHOR	Kendall, John S.
TITLE	A Technical Guide for Revising or Developing Standards and Benchmarks.
INSTITUTION	Mid-Continent Research for Education and Learning, Aurora, CO.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	2001-06-29
NOTE	30p.
CONTRACT	ED-01-CO-0006
AVAILABLE FROM	Mid-continent Research for Education and Learning, 2550 S. Parker Road, Suite 500, Aurora, CO 80014-1678. Tel: 303-337-0990; Fax: 303-337-3005; Web site: http://www.mcrel.org.
PUB TYPE	Guides - Non-Classroom (055)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Benchmarking; Educational Quality; Elementary Secondary Education; *School Districts; *Standards
IDENTIFIERS	Standard Setting

ABSTRACT

This guide describes a process that can be used or adapted by school districts to evaluate standards and guide practitioners in developing locally appropriate benchmarks. The guide is not a training manual that provides a step-by-step guide for undertaking this effort with a group of educators; instead, it suggests an approach and rationale for a review of standards and addresses techniques that might be used when part of the process is adapted for group work at the district level. A content standard is a summary description regarding what it is that students should know and be able to do within a particular discipline. A benchmark is a clear, specific description of knowledge or skill students should acquire by a particular point in their schooling. Performance standards specify how good is "good enough." The guide discusses all of these ways of measuring performance in the following chapters: (1) "Introduction"; (2) "Definition of Terms"; (3) "Review Purpose and Criteria"; (4) "Reference and Comparison Documents"; (5) "Who Will Do the Work?"; (6) "Standards and Benchmark Analysis"; (7) "Grade Placement of Content"; (8) "Coverage of Content"; and (9) "The Larger View." (Contains 58 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

EA

ED 457 198

Mid-continent Research for Education and Learning

A Technical Guide for Revising or Developing Standards and Benchmarks

Regional Educational Laboratory
Contract #ED-01-CO-0006
Deliverable #2001-01

Office of Educational Research and Improvement
U.S. Department of Education
Washington, D.C. 20208

prepared by
John S. Kendall

Mid-continent Research for Education and Learning
2500 S. Parker Road, Suite 500
Aurora, CO 80014
303-337-0990 (phone)
303-337-3005 (fax)
www.mcrel.org

June 29, 2001

TM033296

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

L. Cicchinelli

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

BEST COPY AVAILABLE

© 2001 McREL

To order copies of *A Technical Guide for Revising or Developing Standards and Benchmarks,* contact McREL:

Mid-continent Research for Education and Learning
2550 South Parker Road, Suite 500
Aurora, CO 80014-1678
tel: 303-337-0990
fax: 303-337-3005
web site: www.mcrel.org
e-mail: info@mcrel.org

The compilation and dissemination of this report is sponsored by the U.S. Department of Education, Office of Educational Research and Improvement, under contract number ED-01-CO-0006. The content of this publication does not necessarily reflect the views of any individual laboratory, the Department of Education, or any other agency of the U.S. Government. The publication is available from the education research laboratories that participated in this study.

TABLE OF CONTENTS

Introduction 1

Definition of Terms 1
- Content Standard 2
- Benchmark 2
- Strand/Topic 3
- Performance Standard 3
- Lifelong-Learning Standard 4

Review Purpose and Criteria 4
- Organization 5
- Content 6
- Clarity and Specificity 6

Reference and Comparison Documents 6
- Reference Document 7
- Use of Comparison Documents 7
- Guidelines for Document Selection 8

Who Will do the Work? 11

Standards and Benchmark Analysis 11
- The Language of Benchmarks 12
- Beginning the Review 13
- Benchmark Analysis: An Illustration 14

Grade Placement of Content 16
- Grade Placement: Internal 17
- Grade Placement: External 17
- Grade Placement: The Single "Benchmark" Document 19

Coverage of Content 19

The Larger View 20

Bibliography 22

INTRODUCTION

Although every state but one in the U.S. has state-level standards, some states have been less deeply engaged in standards reform than leading-edge states across the country. Often, those states that are slower to adopt standards also take a decentralized approach to standards in which the responsibility for development of curriculum, instruction, and testing resides at the local level (Laboratory Network Program, 1998). Unfortunately, in some states or local education agencies there is not enough time or resources available to develop standards-based curricula and classroom practices. A number of states do not have curriculum frameworks or grade-level benchmarks to guide practitioners, a critical first step.

Research and experience tell us that major changes in classroom practices can occur when teachers engage in developing standards and grade-level benchmarks and aligning them with high-quality curricula and instructional materials. The purpose of this guide is to describe a process that can be used or adapted by school districts to evaluate standards and to guide practitioners in developing locally appropriate grade-level benchmarks. Curriculum directors and others who might facilitate such a process should find this guide helpful for understanding the underlying technical issues. The guide is not a training manual, however, in that its purpose is not to provide a step-by-step how-to guide for undertaking this effort with a group of educators. Rather, it suggests an approach and rationale for a review of standards, and, where appropriate, addresses those techniques that might be appropriate when a part of that process is adapted for group work at the district level. Later additions to this work will address alternate approaches to a number of techniques presented here, with attention to the advantages and disadvantages of those alternate approaches.

The rudimentary stages of the process described here emerged during McREL's work with a number of school districts that pioneered local standards development in the early 1990s. The approach has been steadily refined and elaborated since that time. In some form or another, the process has been used to revise or develop standards for nine state departments of education and more than 60 school districts across the U.S., as well as education agencies abroad. Although the method is never applied in exactly the same way, it is informed by a set of beliefs and a rationale that guide the decisions that are made in each case; thus it is possible to characterize the work overall as a common process. This should not be considered a "McREL process," but rather, a process commonly used at McREL and one that has served a significant number of clients well over the last ten years.

DEFINITION OF TERMS

Before describing the standards review process, some basic distinctions and concepts should be made clear. These descriptions are definitive for the purpose of this guide only. Although many states and subject-area agencies share these terms and their use, still others adopt different definitions for the same terms or use different terms to communicate very similar ideas. What is common to every endeavor in standards work is the need to make clear the distinctions between types of standards and closely related concepts. The terms to be defined are *content standard*, *benchmark*, *strand* or *topic*, *performance standard*, and *lifelong-learning standard*.

Content Standard

A content standard is a summary description regarding what it is that students should know and/or be able to do within a particular discipline. For example, a standard might state that the student "understands and applies basic and advanced properties of the concepts of geometry."

Content standards primarily serve to organize an academic subject domain through a manageable number of generally stated goals for student learning. These statements help to clarify the broad goals within the discipline and provide a means for readers to navigate the standards document when searching for specific content. The more broadly a standard is described, the more content can be organized beneath it and, thus, the fewer the number of standards needed to encompass the discipline. Conversely, narrowly written content standards are less effective at segmenting a discipline into a manageable number of broad categories, and more of them must be written to address all of the discipline.

An analysis of the content organization of state standards conducted by McREL (Marzano & Kendall, 1996) and more recently by the Council of Chief State School Officers (1998) shows very similar findings. As well as content standards, such labels as "goals," "expectations," and "learning results" serve to identify a similar level of content organization. The number of statements used to organize a discipline may vary — anywhere from 6 to 18, as seen in the CCSSO study of mathematics. Although the language used to construct the statements may vary from one state to another, these standards serve the function of organizing a discipline under the central categories, or organizing ideas, of the discipline.

Benchmark

A benchmark is a clear, specific description of knowledge or skill that students should acquire by a particular point in their schooling. "Students understand basic properties of figures (e.g., two- or three-dimensionality, symmetry, number faces, type of angle)" is an example of benchmark that might be found within a geometry standard.

A benchmark is organized beneath the standard whose content it addresses more specifically — for the benchmark just described, geometry is the standard. A benchmark is assigned to a particular grade level or range of grades. Ideally, a benchmark is placed at the grade at which a student is not only developmentally ready to acquire the understanding or skill it describes, but also at the point in time at which the student has received all prior instruction necessary to learn the new material. Said differently, a benchmark is a grade-appropriate or developmentally appropriate expression of knowledge or skill that is more broadly stated in the content standard.

The specificity of content description identified here by the term "benchmark" appears in other documents under various other names, such as "indicator," "learning expectation," and even "performance standard." As we will see, in some documents, these alternate terms for benchmark might be operationally synonymous with the benchmark, but they may also refer to the fact that in some models of content description, this level of specificity includes a description of student activity. For the purpose of our discussion, we will use "benchmark," or "grade-level benchmark," throughout this guide to indicate a specific description of student knowledge or skill.

The "grain size" of a benchmark cannot be described in absolute terms. One way to consider the level of specificity of a benchmark is that it describes content not so narrowly that it could be mastered by a student in an afternoon, but not so broadly that it might take several months of instruction. Both the finer description of the content (that is, what the student might best learn in the course of a few hours) and the broader description of content (that is, a unit or syllabus that encompasses several months) is best reserved for the curriculum framework, textbook, or similar materials. The benchmark should be specific enough that readers are clear about the instruction and learning it should entail, but not so narrow as to prescribe the day-to-day curriculum nor so broad that the knowledge and skills it describes could be open to numerous equally valid interpretations.

Strand/Topic

A strand (or topic) is a level of content organization that mediates between a standard and a benchmark. Under the geometry standard, for example, topics or strands might include Shapes & Figures, Lines & Angles, or Transformations/Motion Geometry.

Oftentimes a standard is considered too broad an organizational tool for convenient use in lesson planning, reporting, or record keeping. In order to develop a more useful level for working with the content, standards documents sometimes include an intervening level of organization within a document. (For a further discussion of the uses and development of topics, see *Topics: A Roadmap to Standards* [Kendall, 2000]).

One of the reasons that a content standard is so named is to help distinguish standards that describe the content of standards — that is, information and skills — from students' performance, or the demonstration of how well the information or skill has been acquired.

Performance Standard

Performance standards specify 'how good is good enough.' They relate to issues of assessment that gauge the degree to which content standards have been attained." The National Education Standards and Improvement Council (1993, p.iii).

A performance standard describes levels of student performance in respect to the knowledge or skill described in a single benchmark or a set of closely related benchmarks. A performance standard might be described by means of a rubric or a cut-score, or could even be expressed as a percentage correct of the test items designed to assess students on a particular benchmark. The issues that must be addressed in the development of performance standards are complex and beyond the scope of this guide. The benchmark development process described here, however, begins from the assumption that we must first be clear about what students should know and be able to do before we resolve just how well they can understand and apply this information. Put another way, the development of performance standards requires judgment regarding levels of performance, in addition to a judgment concerning what skill and knowledge should be assessed. A process that begins from content standards and benchmarks seeks to address one difficult question at a time.

Lifelong-Learning Standard

A lifelong-learning standard is a summary description regarding what students should know and/or be able to do across a variety of disciplines — for example, "The student applies decision-making techniques."

In addition to discipline-specific knowledge and skills, some skills are not strictly content related but are found in all aspects of the curriculum. These are often called lifelong-learning standards. Lifelong-learning standards may address self-regulation, the ability to work with others, and critical thinking. Although they are "content free" in description, this is because they are and can be applied to content across the curriculum.

Distinctions can be drawn among the terms just defined by considering the purpose and specificity of each. How these distinctions impact the work of standards development and evaluation should become more clear as the process is described in detail. Briefly, it can be said that the purpose of a content standard is to organize subject-area material effectively and provide summary statements that communicate the breadth of a discipline succinctly. The purpose of a benchmark is to state clearly and specifically what the student should know and be able to do at a particular point in schooling. The performance standard describes to what level of detail or with what degree of facility a student has mastered the information or skill identified in a benchmark. Thus, a performance standard does not serve to organize content, nor to identify important information and skills, but rather to characterize performance expectations for students in terms of how well they demonstrate or apply the content described in standards and benchmarks. A lifelong-learning standard, such as one that addresses thinking and reasoning or self-regulation, summarizes a related set of knowledge or skills of a type that happen not to be associated with a single discipline.

REVIEW PURPOSE AND CRITERIA

In order to determine how best to undertake the review or development of standards, those who are engaged in the process should be clear about the purpose or goal of the work. Knowing the purpose of the work will inform decisions regarding what documents to use and how to use them, what methods to employ, and who should be involved in the work.

What emerges as a significant or driving purpose of the work can vary significantly from one school district to another. For example, one district may determine that its standards should be as challenging as those found in countries that have been rated highest on international assessments, while another may want topics or strands to be developed as an additional level of content organization. Among the many clients for whom McREL has conducted standards development or revision, the following set of concerns appear to arise most often:

- Locally developed standards should address the content identified in the state standards.
- The content in the standards document should provide content at each grade level (rather than grade bands).
- There should be no redundant content in standards.
- Students should have an opportunity to learn information and skills prior to being assessed on them.

Less pressing, but still important, are considerations of clarity of language, specificity of detail, and overall coherence or organization of the document.

Purposes can and do vary from one district to another, although many are shared. The following list outlines the criteria used by McREL during the standards review process. The criteria are organized to address the overall organization of content, the content itself, and the clarity and specificity by which the content is communicated.

ORGANIZATION:

- Do the standards work as organizing statements of the discipline?

  Do the set of standards preclude the problem of having benchmarks appear under more than one standard, owing to overlapping statements?

  Are the standards organized hierarchically? That is, are more specific benchmarks organized under more general standards?

  Is it clear to the reader what content will be found organized beneath each standard?

  Will the content organization facilitate the construction of lesson plans and units?

  Will it facilitate grading and reporting?

- Are the benchmarks organized appropriately?

  Do benchmarks appear under the appropriate standard?

  Does the same benchmark appear under more than one standard?

  Are the standards and benchmarks useful for grading and reporting purposes?

CONTENT:

- Are the important knowledge and skills of the discipline addressed?

  Are the significant knowledge and skills of the discipline addressed?

  Are the knowledge and skills identified in the state standards appropriately reflected?

  Are the knowledge and skills that are assessed by the district and/or state appropriately reflected?

- Do the knowledge and skills appear at the appropriate level?

  Are the benchmarks appropriately "scaffolded" so that, where possible, logically prior knowledge and skills appear in the correct sequence?

  According to the best information available, are the knowledge and skills placed at levels appropriate to the developmental level of the student? Are the benchmarks appropriately challenging?

  Given the level at which knowledge and skills are assessed (via standardized or other assessments), do the knowledge and skills appear at an early enough grade level such that students have adequate opportunity to learn them prior to being tested?

  Does the content, taken as a whole, represent what is manageable for instruction, given the time available in the school day?

CLARITY AND SPECIFICITY:

- Is the language clear and free of jargon? If there are technical terms, are they defined?

- Is the language specific enough that all stakeholders know what it is they will be held accountable for?

- Is it clear in the document what material is presented as an example as opposed to what material students must learn?

- Is it clear what is expected of students by the end of each grade?

Once the central purpose of the work is made clear, the documents that will be consulted in the work can be considered.

REFERENCE AND COMPARISON DOCUMENTS

It is useful in the evaluation or development of standards to think of the work as involving a single reference document and a set of comparison documents.

Reference Document

The reference document is the document from which the work begins. As the basis for the development of a district or school standards document, the reference document should possess characteristics and qualities that accord well with the purposes that have been established for the standards development effort from the outset. For example, if the principal reason for the development of standards is that the district is required to address standards that have been promulgated by the state, then the state standards document would, in most cases, be the logical starting point. If, however, the state document is known to be vague or unclear, and it appears that another, highly regarded document can accommodate the content addressed in the state standards, yet also presents material clearly and concretely, then the state standards document may not be the best choice as a reference document. In such a case, the state document would be used as a comparison document, so that it can be consulted to determine whether, in fact, all the significant content addressed in the state document is addressed in the final standards document.

Use of Comparison Documents

Just as the reference document should be selected in consideration of the purpose determined at the beginning of the process, so should the comparison documents. It is convenient to categorize comparison documents by the purposes they are intended to serve, although, of course, any single document may serve multiple purposes. Comparison documents may be useful to help evaluate the

- appropriate breadth of coverage,
- depth of coverage,
- appropriate level of difficulty,
- level of specificity (including examples),
- grade placement, and
- content selection /reduction.

Additional documents may be used at a more general level. For example, some documents might be consulted as models of content description or models of standards structure, rather than for specifics of content. Such documents could help inform

- models for standard structure (including the use of strands), and
- clarity of design (e.g., the use of introductory sections, or design features).

Depending upon the purpose for the review, the comparison documents might include

- standardized assessments,
- local assessments,
- standards documents from other states,
- standards documents from other countries,
- standards documents from national subject-area organizations, and/or
- standards documents from other agencies and organizations.

Guidelines for Document Selection

In its work, McREL relies upon comparison documents in order to evaluate and revise content, and that is the approach that will be the focus of this guide. The approach offers advantages over relying upon one or two content experts to undertake the evaluation of a document, a technique used by a number of organizations, such as the CCSSO, that have conducted standards reviews in the past. Our concern is that a single expert or even a small panel could possess an ideological view that is not shared by the larger community and, more significantly, by the community of teachers and stakeholders whose standards are under review. Such a view might become obvious, as when reviewers for the Fordham Foundation award poor marks to those mathematics standards that encourage the early use of calculators among students and who "deplore" the "enthusiasm" they perceive on the part of NCTM in endorsing such a view (Raimi & Braden, 1998). Unless such statements are overtly made in the course of a review, one cannot be certain what biases might inform the expert critique. If, instead, carefully selected comparison documents are used to inform the evaluation, it is less likely that an ideological — or idiosyncratic — view has affected the process. There are three requirements that a reference and comparison document should meet, in order to insure good and useful data for comparison:

1. The document was constructed through a consensus process that included experienced teachers and educators at all levels and incorporates best available research regarding appropriate grade-placement of content.

2. The document has been checked against criteria by those who will use it to conduct the review.

3. The document has been rated highly by more than one of those organizations that has undertaken a nationwide review of state documents.

The first criterion guards against the use of a document that does not reflect a view commonly shared by educators of the subject area or is uninformed by research. The criterion also increases the likelihood that the document reflects the collective wisdom of teachers and their understanding about what students are capable of and at which grades.

The second criterion simply requires that a comparison document be evaluated to determine that it is appropriate to the task, that is, that the document fares well on the criteria for which it is being used to evaluate the reference document. For example, a given document might not be the best example of how information and skills should be described at the benchmark level, but if its primary use is to help revise the content organization of the document — the revision of standards, and the distribution of benchmarks — it should of course be highly valued on that criterion.

The third criterion, which is applicable only to state standards documents, seeks to determine the relative value of a document through its ranking by organizations that have rated the state standards documents. Three such organizations have published critical reviews: the American Federation of Teachers, the Council for Basic Education, and the Fordham Foundation. Unfortunately, as has been noted (Olsen, 1998), the views of these groups can vary significantly. It is possible, however, to select state standards documents that have been generally well-rated by all organizations, which should offset the apparent conflicting criteria, or application of that

criteria, that the disparate ratings reflect. McREL has conducted a review of the ratings provided by these national organizations in order to determine the five states that are most highly rated within each subject area for their coverage of the subject matter and the clarity of their presentation. The following state standards documents, which are organized by content area, should prove useful for consultation during the review and revision of a standards document.

English Language Arts:

English-Language Arts Content Standards for California Public Schools, Kindergarten Through Grade Twelve (1998), by the California Department of Education

The English Language Arts Curriculum Framework (1997, February), by the Massachusetts Department of Education

Language Arts Standards (1999), by the Arizona Department of Education

Standards of Learning for Virginia Public Schools: English Standards of Learning (1995, June), by the Board of Education, Commonwealth of Virginia

Wisconsin's Model Academic Standards for English Language Arts (1999), Wisconsin Department of Public Instruction

Mathematics:

The California Mathematics Academic Content Standards (Prepublication Ed.) (1998, February 2), by the California Department of Education

Core Curriculum Standards: Mathematics (1994), by the Utah State Office of Education

Model Competency-Based Mathematics Program (1990, November), by the Ohio Department of Education, Division of Elementary and Secondary Education

Standards of Learning for Virginia Public Schools: Mathematics Standards of Learning for Virginia Public Schools (1995, June), by the Board of Education, Commonwealth of Virginia

West Virginia Programs of Study: Instructional Goals and Objectives (1995, June), by the West Virginia Department of Education

Science:

Rhode Island Science Framework (1996, August 14), by the Rhode Island Department of Education

Science Content Standards Grades K–12 (Prepublication Ed.) (1998, February 2), by the California Department of Education

Science Curriculum Framework (1998, March), by the Connecticut State Department of Education, Division of Teaching and Learning

Science Language Arts Curriculum Framework (1995, June), by the Delaware Department of Education

Science Standards (1998, August 24), by the Arizona Department of Education

Social Studies:

Geography

Alabama Course of Study: Social Studies (1998, February), by the Alabama State Department of Education

Curriculum Standards: Social Studies (2000), by the South Carolina State Department of Education

Kansas Curricular Standards for Civics-Government, Economics, Geography, and History (1999, July), by the Kansas State Board of Education

Social Studies Content Standards (1997, May), by the Louisiana State Department of Education

Social Studies Standards (2000), by the Arizona Department of Education

History

Alabama Course of Study: Social Studies (1998, February), by the Alabama State Department of Education

History-Social Science Content Standards for California Public Schools: Kindergarten Through Grade Twelve (2000), by the California Department of Education

Kansas Curricular Standards for Civics-Government, Economics, Geography, and History (1999, July), by the Kansas State Board of Education

Social Studies Standards (2000), by the Arizona Department of Education

Standards of Learning for Virginia Public Schools: The History and Social Science Standards of Learning (1995, June), by the Board of Education, Commonwealth of Virginia

Researchers at McREL have conducted a study of these documents in the language arts, mathematics, and science (Kendall, Snyder, Schintgen, Wahlquist, & Marzano, 1999) and geography and history (Kendall, Schoch-Roberts, & Young-Reynolds, 2000). These reports provide a listing by subject area of the content that was found to be common among these highly

rated standards documents. As will be noted in the Content Coverage section, these reports could prove useful when districts seek to determine what content is most essential.

The individual state standards documents should be particularly helpful when outside comparison is needed to determine at what grade a benchmark would best be placed. Absent a substantive body of knowledge based on empirical research on such a question, we must rely heavily upon the knowledge and experience of classroom teachers and other educators to help inform decisions concerning the appropriate grade placement. Considering the consensus process by which they were formed, these documents represent that information fairly well.

WHO WILL DO THE WORK?

Having determined the purpose of the work and selected the necessary documents, the school district should next determine how the work will be undertaken. There are two common ways in which school districts undertake a standards review. One is to draft a set of standards, using teacher expertise with the guidance of a curriculum director and/or someone trained in a benchmarking process. Typically, and ideally, this work is then submitted to an outside organization, such as McREL, which can provide an objective comparison against selected documents and according to the criteria established in conference with the school district.

We have found that there are two significant difficulties with this approach. First, in small school districts there may not be subject-area specialists available to help guide the work. When individuals are available, however, the chief difficulty with this scenario is that, although teachers gain considerably in their understanding of the rationale for standards, and have the opportunity to become familiar with the content of standards, the work itself is particularly taxing and can wear the patience of even the most resolute among them. Beyond a certain point in the process, the exercise may become tedious and, consequently, the quality of the work may suffer. The quality can be redressed through the review of an outside organization, but the work itself may sap enthusiasm for the standards enterprise. In short, although major changes in classroom practices can occur when teachers engage in developing standards and grade-level benchmarks, such work must be carefully modulated so that the process itself does not become burdensome.

Another common approach is to have the outside organization create the first draft of standards based on the district's criteria. This draft is submitted to a panel of district teachers, who, with the guidance of a curriculum director and/or others trained in the process, review and refine the document in a way that reflects their particular concerns. This approach is less demanding for teachers but provides them with an opportunity to engage in the process and contribute their expertise. Once the teachers have reviewed and/or revised the draft, the document is resubmitted to the external organization, which typically incorporates the requested changes, while noting any reservations concerning content that has been changed.

STANDARDS AND BENCHMARK ANALYSIS

Whether the initial draft is undertaken by the district or by another organization, the review commences once the primary purpose and other significant criteria for the task have been

determined, the comparison documents selected, and those who will take on the work have been convened and briefed on the task.

The work itself requires a deliberate choice about the language that will be used to describe and organize content. It is also important to know where to begin the analysis.

The Language of Benchmarks

As noted earlier, two critical distinctions between a content standard and a benchmark center on their differing purposes and level of specificity. Because a content standard is used primarily to organize content, the language of the standard is less critical than the topic it identifies. For example, although the wording differs considerably among them, any of the following state standards successfully demarcate the area of geometry within mathematics:

> **Colorado**: Students use geometric concepts, properties, and relationships in problem-solving situations and communicate the reasoning used in solving these problems.
>
> **Kansas**: The student uses geometrical concepts and procedures in a variety of situations.
>
> **Missouri**: Geometric and spatial sense involving measurement, trigonometry, and similarity and transformations of shapes
>
> **North Dakota**: Students understand and apply geometric concepts and spatial relationships to represent and solve problems in mathematical and nonmathematical situations.
>
> **South Dakota**: Students will use the language of geometry to discover, analyze, and communicate geometric concepts, properties, and relationships.
>
> **Wyoming**: Students apply geometric concepts, properties, and relationships in problem-solving situations. Students communicate the reasoning used in solving these problems.

This level of generality, although useful for the standard, should not characterize the benchmark. Because the benchmark should communicate clearly what students should know and be able to do, the language must be more concrete and precise. There are a number of ways in which benchmarks — or their equivalents — are communicated in various standards documents. Some describe content in terms of a simple student activity; others might describe a performance task — that is, an extended task that includes the context within which the student acquires and demonstrates knowledge.

We have found that the clearest level of content description, and the one best suited for the evaluation of content, maintains a distinction between two types of knowledge: information and skills. There are several reasons for this. First, the information and skills description, as opposed to the activity or task description, does not require the reader to make inferences from the activity or task to the information and skills that would be required for successful demonstration

of that task; rather, student information and skills are described in a straightforward manner. Second, the activity or task description tends to be narrowly prescriptive in that it characterizes not only what the student should know and be able to do, but how the student should demonstrate this knowledge. Thus, the content described is likewise narrowed; users might erroneously believe that the information or skill required by the activity or task is a complete description of the information or skill the student should acquire. Finally, although a task or activity might be useful for teachers as a guide for instruction or classroom assessment, it is not useful for teachers as a guide to what information and skills are essential for students to learn. Such activity descriptions confound the issue of how students are to demonstrate competence with the logically prior and equally significant issue of what the content of the curriculum should comprise. Once the content has been determined, of course, delineating various ways in which knowledge might be presented and demonstrated is appropriate. Until and unless that content is identified, however, we believe it is best to keep the two kinds of description separate. This view is based on the same rationale, described above, that argues for the establishment of content standards and benchmarks prior to the development of performance standards.

As a consequence of the decision to consider the content of benchmarks apart from descriptions of student activity or performance, the method of standards review that we use has some noteworthy characteristics. Specifically, at the benchmark level we analyze material to determine what information the student should know and, separately, what skills a student should acquire. Such distinctions have proved useful in theories of learning and cognition (Anderson, 1993; Keil, 1989; Damasio, 1994). One type of knowledge relates only to information, not skill. Acquiring this type of knowledge involves understanding its component parts. For example, knowledge of the concept of "a geographic region" includes understanding the characteristics of a variety of regions, knowing criteria that give a region identity, understanding how regional boundaries can change, and so on. This type of knowledge is commonly called *declarative knowledge*. One might think of such knowledge as composed of the *information* important to a given content area. Information includes such things as facts, events, episodes, concepts, principles, and generalizations.

Another type of knowledge, *procedural knowledge*, can be thought of in terms of skills or processes. A process may or may not be performed in a linear fashion. For example, performing long division is a process: you perform one step, then another, and so on. Reading a map also involves certain steps, but these steps, unlike those in long division, do not have to be performed in any set order. You might read the name of the map first, then look at the legend, or you might just as effectively perform these steps in reverse order. Some skills, however, like algorithms, require adherence to a particular sequence. One might think of procedural knowledge as the skills and processes important to a given content area. It is noteworthy that a recent review of the content areas (Kendall & Marzano, 2000) supports a commonly held belief that the subject areas of language arts and mathematics contain a relatively high proportion of procedural knowledge as opposed to declarative knowledge.

Beginning the Review

Early in McREL's standards work, McREL researchers often developed content standards at the end of the development process, that is, after the information and skills had been identified and described as benchmarks. Benchmarks addressing similar ideas were then clustered together and the clusters were organized until fairly robust categories were formed. This work "from the

bottom up" was somewhat laborious; there were numerous false starts, and occasionally a problem in the organizational scheme did not become apparent until many of the items had been sorted beneath standards. Such work still might be necessary when new content areas are developed.

Many standards documents have been produced since that time, however, offering a number of different ways to organize content. Because the task is made considerably simpler when the review work is conducted standard by standard — or even more specifically, by a topic or strand within a standard — we recommend that the review begin not by the analysis of any statements of knowledge or skill that might be at hand, but first by the selection of a standard, that is, a category within which to start the work. The standard should be straightforward, that is the description of what it comprises should be unambiguous. The standard might be compared to those found across states to determine if it is common. If it is not a common standard, then it is better addressed once the process has become familiar.

Benchmark Analysis: An Illustration

An appraisal of the standards themselves as categories can take place in the course of the benchmark review. Reviewers should simply note whether each benchmark fits logically within the standard in which it is placed. If a benchmark could be categorized under more than one standard (after the benchmark has been analyzed and revised), then the standards will likely need to be re-formed so that they do not overlap the same content. It is best, especially for those new to the work, to begin selecting benchmarks from within a standard that appears to be fairly clearly defined.

Within the selected standard, the work commences with a benchmark. The first question, and one that will continually be repeated in the review process, is, "What knowledge and skill is communicated in the benchmark?" If the reference document under review distinguishes between declarative and procedural knowledge, the answer to this question could be quite straightforward. Often, however, benchmarks are written in such a way that the content is not easily deciphered. Consider, for example, the following:

> *Students explore and develop relationships among two- and three-dimensional geometric shapes.*
>
> *New York State Standards – Mathematics, Elementary*

This statement does not describe the knowledge a student should have, nor a skill; rather, it describes an activity in which the student might engage. It is in fact difficult to determine for what purpose the activity was designed, or what the student will know and be able to do as the result of having explored shapes or "developed relationships" among them. In the course of reviewing a benchmark, a useful analytic question when confronted with an activity is, "What knowledge or skill should a student acquire or have acquired in order to engage in this activity successfully?" Although the benchmark noted above resists that kind of inquiry, such a question can be productive when considering an activity such as the following:

[Students] draw two- and three-dimensional geometric shapes and construct rectangles, squares and triangles on the geoboard and on graph paper satisfying specific criteria.

Pennsylvania State Standards – Mathematics, Grades K–3

Here we can infer that the student who successfully manages the activity knows common two- and three-dimensional geometric shapes including rectangles, squares, and triangles and understands that shapes fit certain criteria in order to be named as they are. It is important to note that although the benchmark is introduced with the verb "draw," it is not the skill of drawing that is the focus anymore than is the ability to use a geoboard. Thus, we know that this is fundamentally about information, that is, it is a declarative benchmark, not a procedural benchmark, because it does not concern itself with a particular skill. That the information is introduced via the activity of drawing might confuse some readers or at the least make the content of the benchmark less clear than it could be.

The initial benchmark of this exercise remains problematic, however, because it simply directs that students explore shapes and "develop" the relationships among them. Because we cannot deduce knowledge and skill from a description such as this, we adopt two strategies to find a solution. One is to examine related content at the grade band prior to and the grade band following the item within the same document; the other is to consult the comparison documents. In the case of our example, elementary is the earliest grade band in the document, so we have no prior grade band to consult. At the next grade band, we find this more concrete information:

Students use properties of polygons to classify them.

New York State Standards – Mathematics, Intermediate

By this we understand that students at the intermediate level should have advanced enough to discriminate among polygons based on their properties. It seems probable, then, that prior to this time, students should have understood distinctions between polygons and non-polygons, for example, circles. If we refer to the McREL study (Kendall et al., 1999) as a guide, we see that among highly rated state standards documents in mathematics, the following content related to geometry for the grades 3–5 is typically found:

[The student] understands basic properties of figures (e.g., two- or three-dimensionality, symmetry, number of faces, type of angle).

[The student] knows basic geometric language for describing and naming shapes (e.g., trapezoid, parallelogram, cube, sphere).

While at grades K–2:

[The student] understands the common language of spatial sense (e.g., "inside," "between," "above," "below," "behind").

We now have some concrete information, which, along with the content at the intermediate level (usually interpreted as grades 3-5) of the reference document, will help to revise the material. The first stage of the process, then, is to determine as well as possible what knowledge or skill is the focus of the benchmark and to revise the language until that focus is clear. In order to do this, it is sometimes necessary to read the benchmarks within the standard at other grades or grade bands within the reference document. On occasion, a benchmark will contain within it both declarative and procedural knowledge. In keeping with this model of content description, such

benchmarks should be rewritten to make the differences clear. For example, consider the following benchmark:

> *Students explore and express relationships using variables and open sentences.*
> *New York State Standards – Mathematics, Elementary*

A review of the elementary level for this standards document makes clear that students have not yet been formally introduced to the concept of a variable, which is declarative information. Consulting other documents, such as *Principles and Standards for School Mathematics* (NCTM, 2000), we confirm that it is appropriate for students to understand the concept of variable at the 3–5 grade band, which appears to correspond with the New York State standards elementary level. The benchmark should be divided into two separate benchmarks to distinguish between declarative and procedural information, for example:

> *Students know that a variable is a letter or symbol that stands for one or more number.*
>
> *Students use open sentences to represent problem situations.*

Although this process brings clarity and specificity to the benchmarks, we still have not determined at what grade to assign the content.

GRADE PLACEMENT OF CONTENT

A commonly expressed goal for many school districts as they revise a standards document is to develop grade-by-grade benchmarks from their state's standards document, which is commonly written at grade bands, such as K–4, 5–8, and 9–12. In order to do this work, the benchmark found at a grade band must first be analyzed for clarity of content and revised, if necessary, before it is placed at the appropriate grade level.

Few have the expertise or the time to examine all research — scarce as that might be — that could be brought to bear on the appropriate grade placement for each of the hundreds of benchmarks that form a discipline. Although documents produced by national subject-area organizations might hold the greatest authority in this regard, none provides grade-by-grade recommendations. It seems probable that these organizations were deterred by the lack of research supporting placement of content at a particular grade level. In addition, many organizations avoid the assignment of content to a grade because it could be perceived, understandably enough, as overly prescriptive. Yet, the problem for a school district remains. Content must be assigned to a grade because it must be taught at a grade.

For convenience of discussion, the factors that influence or inform grade placement can be classified as internal and external. From "within," the placement of content might depend upon the student's developmental or psychological readiness to learn the content, the logical progression of content within a topic, or, when neither case can be made, the grade at which the content has been traditionally placed. An example of the latter kind of decision is typical in the social studies. For example, there are district curricula that address early hunter-gatherers in first grade; others that do not address such a topic until 6th grade or later. External factors are those factors that affect the grade placement of content because of pressures brought to bear either by assessments or a desire to be viewed positively in comparison with other schools.

Grade Placement: Internal

Because there is no definitive work available on what knowledge and skills should be addressed at each and every grade level, the work of grade placement of content is at best an "educated guess," at worst, an arbitrary assignment. Because there is no nationally developed standards document that recommends particular grade levels for content, we turn to those highly rated state standards documents to provide a consensus view of experts in the field. A single state standards document may be too limited a resource. McREL's experience suggests that consulting two grade-by-grade documents and, if time and resources permit, three such highly rated documents, provides a wealth of useful information.

For example, in the preceding section, a mathematics benchmark appearing at the elementary level (in New York State, roughly corresponding to grades K through 5) was rewritten to distinguish declarative knowledge regarding the concept of a variable. A review of comparison documents reveals the following information regarding the grade placement for the concept of a variable:

> *The student will investigate and describe the concept of variable.*
>
> *Virginia State Standards – Mathematics, 5th grade*

> *[The student] knows that a variable can be used as a placeholder for a specific unknown*
>
> *Utah State Standards – Mathematics, 5th Grade*

> *Students use letters, boxes, or other symbols to stand for any number in simple expressions or equations (e.g., demonstrate an understanding and the use of the concept of a variable).*
>
> *California State Standards – Mathematics, 4th Grade*

This information, at least by a preponderance of the evidence, suggests that the 5th grade is most appropriate for a benchmark on the concept of a variable. If a portion of this analytic work is done within a district or if teachers review a draft prepared by an external agency, they can contribute to such decisions through their own teaching experience and their knowledge about where the content is currently taught in their school.

Grade Placement: External

Externally, the factors that influence grade placement are primarily state assessments, standardized assessments, and "benchmark" documents. External factors tend to have a greater impact on grade placement than do internal factors. For example, if it can be determined from released forms of a state and/or a standardized assessment that students are tested in the late spring of grade five or the early fall of grade six on the concept of a variable, it is likely that the district would keep the benchmark at grade 5, where most state standards documents have placed it. If however, the concept is assessed at spring of 4th grade or early fall of 5th grade, districts, in our experience, will request that the benchmark be set to grade four.

Certainly, earlier forms of a state or standardized assessment, if available, can be used to inform grade placement. There are a number of caveats to this approach, however. First, it is important to remember that many standardized assessments are not designed to measure students against a

criterion, but in relation to their peers; that is, they are norm referenced. This fact can result in unintended consequences when assessment items are used to determine grade placement of content. For example, it is commonplace in norm-referenced testing to place a few items at a difficulty level so that many, if not most students, will not successfully answer these items. Such an approach permits the ranking of those students who are at the upper end of the scale. For example, on a test designed for fourth graders, some items will actually be appropriate for fifth, or even sixth graders – only a handful of fourth graders will answer them correctly. These items are included in the assessment precisely because most student *should not* be able to answer them, not because the test developers have determined to reform the curriculum via the assessment. If those who use norm-referenced assessment are not aware of these out-of-level items, they could understandably seek to address all content that is tested, even if other source material suggests that it is inappropriate for the grade. If every item on a standardized test were accepted as the norm for the grade level it tests and used to place benchmarks accordingly, one could imagine a race that never ends. First, those who review the standards would move a benchmark to a lower grade because the content was found to be assessed at the lower grade on a norm-referenced test. As students acquire the knowledge and skill in the benchmark, they learn to master the once difficult test item; assessment developers in turn discover that the item no longer helps to sort out students at the upper end, so they add a still more difficult benchmark. Of course, this scenario assumes that students can master the material simply because it is a part of instruction, when in fact, the material is only present because the test developers believed that most would not master it. Clearly, a simple acceptance of item placement on a norm-referenced test without reference to other documents is not an informed method for adjusting content for grade level.

Another problem is the tendency to "overinterpret" a test item, that is, to conclude that a test item dictates that a benchmark should be placed at an earlier grade when, in fact, the argument cannot be supported. In some cases, the test item actually measures a more rudimentary skill than the reviewer believes. Again, just as 'above grade' items are deliberately placed in an assessment, so also 'below grade' items help to establish a lower threshold. Another problem in overinterpretation occurs when an item does not directly match the benchmark, but is still accorded enough weight that it changes the grade placement of the benchmark.

We have found two methods that help to mitigate against such overinterpretations. The first is to use two raters who independently review test items against the standards. Discrepant views must be resolved before the test item can be used to determine grade placement. The second technique is to allow raters to characterize an item as a direct match or an indirect match to a benchmark. A direct match means that it appears highly likely that a student who mastered the benchmark would successfully answer the related test item. An indirect match indicates that it is possible, in the course of instruction on a benchmark, that the student might well learn the knowledge and skills that would result in the student's success on the test item, but it is not clear that this is the case. Using this method, both types of matches would be noted next to the benchmark, but only direct matches — agreed-upon by both raters — would cause a benchmark to be moved to another grade level. More sophisticated variants of this approach could be adopted, such as using more than two raters, employing a Likert scale to extend the range of possible ratings on the "goodness of fit" between an item and benchmark, and characterizing the "goodness of fit" on more than a single dimension.

Grade Placement: The Single "Benchmark" Document

While both reference and comparison documents are used to select and organize content, some districts use a single benchmark document to place content at a specific grade level. This is an approach that we counsel against. A benchmark document refers to a document that sets an absolute standard against which the reference document is evaluated. In contrast to a set of comparison documents, a benchmark document dictates to what level of detail various topics within the standards should be addressed and at what grade level benchmarks should appear in order to be considered appropriately rigorous. A state standards documents may be used as either an "internal" *comparison* or "external" *benchmark* document, depending upon how it is used to inform content selection. If the *California Mathematics Academic Content Standards* was chosen as a benchmark document, then we would expect that the concept of the variable in our example would be set at the 4th grade level, effectively overruling the information available from the other documents.

Selecting one document that is used to "trump" all other grade placements gives undue emphasis to a single document. For example, a review of the math documents from Japan (*Mathematics Program in Japan* [1990]) and Singapore (*Primary Mathematics Guide* 1A & 1B [1994]) shows that in these documents as well, the concept of a variable does not appear until grade 5. Yet students in both of these countries performed exceptionally well on an international assessment (the Third International Mathematics and Science Study).

In addition, the use of a single benchmark document does beg a number of questions — for example, whether placing content as it appears in the document of a competitive country gives sufficient regard to the cultural differences, including differences in education systems, that might account for the relative success of students on international assessments.

COVERAGE OF CONTENT

A common concern for many districts is whether all the significant content within a domain is being addressed. It is not unusual for a school district to request that the content in their state's standards document be compared against authoritative documents to ensure that all important knowledge and skills are covered. The goal is laudable and, given the low quality of some state standards documents as others have noted (American Federation of Teachers, 1999; Council for Basic Education, 1998; Fordham Foundation [Finn & Perilli], 2000), probably a good idea.

However, a significant difficulty arises when standards are reviewed for their coverage of a given domain, for a considerable amount of content has been identified as important in each discipline. It has frequently been observed that all of the knowledge and skills identified as important by national organizations cannot be addressed in the classroom given the time available in the school day. Education researcher Chester Finn, after reviewing documents produced by many standards-setting groups, asserted that "the professional associations, without exception, lacked discipline. They all demonstrated gluttonous and imperialistic tendencies" (in Diegmueller, 1995, p. 6). A similar perspective can be found in the report of the Third International Mathematics and Science Study (TIMSS), a large-scale, cross-national comparative study of math and science curricula. In addressing the relatively poor performance of U.S.

students, the report's authors note that our "preoccupation with breadth rather than depth, with quantity rather than quality, probably affects how well U.S. students perform in relation to their counterparts in other countries" (Schmidt et al., 1997, p.2). Researchers Marzano and Kendall (1999) show that, at least by one measure, attempting to address all the content identified in standards documents would mean that "schooling would have to be extended from kindergarten to grade 21" (p. 104).

Two questions define the central problem of content coverage: How much time for learning is available from kindergarten through the 12th grade? and How much time is needed for a student to learn and master the content communicated in a single benchmark? With regard to the first question, a guide for helping to determine the amount of time available for learning, along with an overview of studies devoted to the question can be found in Marzano and Kendall (1999, pp. 99–113). As to the second question, at least one study has been undertaken in an attempt to determine the time required for students to learn and master a benchmark, by interviewing teachers about their perceptions (Florian, 1999). Much more study will be needed to answer this question, however.

Even if these questions are fully answered — if we know how much time is available for teaching K–12 and how much time is needed to learn each and every benchmark — the content coverage difficulty could not be resolved unless there were also some means for effectively reducing the number of standards and benchmarks to that which could be covered. One practice can be adopted as part of the standards review process to help with this selection. During the review, each benchmark can be annotated to reflect the documents in which closely related content was found. At the end of the review, the benchmarks can be ordered to reflect their relative importance as indicated by the frequency with which the knowledge and skill they identify was found in the documents. As part of this calculation, we recommend noting whether the content of a benchmark has been identified as common among the highly rated state standards documents for the subject area. Two studies, referenced earlier (Kendall et al., 1999, 2000), can supply this information. The ranking method can be refined fairly easily — for example, greater weight can be accorded to a district's state standards or assessments than to other documents. Ranking content in such a way may appear too data driven a solution to some; however, it does present a preferable alternative to arbitrarily ignoring content when time is at a premium.

THE LARGER VIEW

To summarize, each benchmark in the document should be analyzed, revised if necessary, placed at the appropriate grade level, and annotated according to which documents support the content described. Ideally the content should be rank ordered to reflect how frequently it is cited in those documents that are consulted during the review.

The benchmarks should also be considered in terms of how they relate to associated benchmarks in the grades preceding and following them. In the demonstration sample, other grades in a standards document were consulted to resolve a problem of interpretation (i.e., regarding how much a student should know about shapes and their characteristics). However, the review of the "vertical alignment" of related content should be conducted formally, not simply as problems arise. This step is intended to ensure that no significant benchmarks have been omitted that are

necessary for mastery of content. We rely on comparison documents and a "preponderance of the evidence" from them about which content should be included prior to a given benchmark, and which afterwards. When districts take on this work, cross-grade teams can help ensure the coherence of the document.

At the same time, a review of the benchmarks should also be conducted "horizontally," that is, within the grade itself, to ensure that the same or very similar content has not been repeated elsewhere. Such can be the result of poorly written standards, which permit very similar content to be organized beneath more than one standard.

Finally, the usefulness of the content organization itself should be considered. Would the document be more useful if another organizing layer of content were added, such as a topic or strand? Would the document be more clear if there were sample activities paired with each benchmark?

The revision or development of a standards document is a complex task, requiring patience, attention to detail, and a clear understanding of the goals and hazards of the enterprise. This overview of the technical issues involved and the processes used at McREL is intended to help others who are about to undertake this work.

BIBLIOGRAPHY

Alabama State Department of Education. (1998, February). *Alabama course of study: Social studies*. Retrieved June 20, 2000 from the World Wide Web: http:// www.alsde.edu/default.asp?info=2&toc=2&sSectionID=8&ProjectID=109

American Federation of Teachers. (1999). *Making standards matter*. Retrieved March 1, 2000, from the World Wide Web: http://www.aft.org/edissues/standards99/Toc.htm

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Arizona Department of Education. (1998, August 24). *Science standards*. Phoenix, AZ: Author. Retrieved June 17, 1999 from the World Wide Web: http://www.ade.state.az.us/standards/science/

Arizona Department of Education. (1999). *Language arts standards*. Phoenix, AZ: Author. Retrieved May 18, 1999 from the World Wide Web: http://www.ade.state.az.us/standards/language-arts/

Arizona Department of Education. (2000). *Social studies standards*. Retrieved June 20, 2000 from the World Wide Web: http://www.ade.state.az.us/standards/sstudies/standard1.htm

Board of Education, Commonwealth of Virginia. (1995, June). *Standards of learning for Virginia public schools: English standards of learning*. Richmond, VA: Author.

Board of Education, Commonwealth of Virginia. (1995, June). Standards of learning for Virginia public schools: Mathematics standards of learning. Richmond, VA: Author.

Board of Education, Commonwealth of Virginia. (1995, June). Standards of learning for Virginia public schools: The history and social science standards of learning. Richmond, VA: Author.

Board of Education, Commonwealth of Virginia. (1995, June). *Standards of learning for Virginia public schools: Science standards of learning*. Richmond, VA: Author.

California Department of Education. (1998). *English-language arts content standards for California Public Schools: Kindergarten through grade twelve*. Sacramento, CA: Author. Retrieved June 14, 1999 from the World Wide Web: http://www.cde.ca.gov/board/standards.html

California Department of Education. (1998, February 2). *Science content standards grades K–12* (prepublication ed.). Sacramento, CA: Author. Retrieved June 17, 1999 from the World Wide Web: http://www.cde.ca.gov/board/science.html

California Department of Education. (1998, February 2). *The California mathematics academic content standards* (prepublication ed.). Sacramento, CA: Author. Retrieved February 1999 from the World Wide Web: http://www.cde.ca.gov/board/k12math_standards.html

California Department of Education. (2000). History-social science content standards for California public schools: Kindergarten through grade twelve. Sacramento, CA: Author.

Colorado Board of Education. (1995). *Colorado model content standard for mathematics.* Retrieved May 20, 2001 from the World Wide Web: http://www.cde.state.co.us/download/pdf/math.pdf

Connecticut State Department of Education, Division of Teaching and Learning. (1998, March). *Science curriculum framework.* Hartford, CT: Author. Retrieved June 17, 1999 from the World Wide Web: http://www.state.ct.us/sde/brta/framewrk/frame.pdf

Council for Basic Education. (1998).Great expectations: Defining and assessing the rigor in state standards for mathematics and English language arts. Washington, DC: Author.

Council of Chief State School Officers. (1998). *Tool kit: Evaluating the development and implementation of standards.* Retrieved May 29, 2001 from the World Wide Web: http://www.ccsso.org/tk98.html

Damasio, A. (1994). *Descartes' error.* New York: Grosset/Putnam.

Delaware Department of Education. (1995, June). *Science language arts curriculum framework.* Dover, DE: Author. Retrieved August 18, 1999 from the World Wide Web: http://www.doe.state.de.us/Standards/Science

Diegmueller, K. (1995, April 12). Running out of steam. Struggling for standards: An Education Week special report. Washington, DC: Education Week.

Finn, C. & Petrelli, M. (Eds.). (2000). *The state of state standards 2000.* The Thomas B. Fordham Foundation. Retrieved July 2, 2000 from the World Wide Web: http://www.edexcellence.net/library/soss2000/ standards%202000 .html#standards

Florian, J. (1999). *Teacher survey of standards-based instruction: Addressing time. Technical report.* Aurora, CO: Mid-continent Research for Education and Learning.

Japan Society of Mathematical Education. (1990, January).*Mathematics program in Japan: Kindergarten to upper secondary school.* Retrieved March 15, 2001 from the World Wide Web:http://carson.enc.org/TIMSS/timss/teaching/125447/5447.htm

Kansas State Board of Education. (1999, July). *Kansas curricular standards for civics-government, economics, geography, and history.* Retrieved June 20, 2000 from the World Wide Web: http://www.ksbe.state.ks.us/outcomes/socialstudies.html

Kansas State Board of Education. (1999, July). *Kansas curricular standards for mathematics.* Retrieved May 20, 2001 from the World Wide Web: http://www.ksde.org/outcomes/math399.html

Kendall, J. S., (2000). Topics: A roadmap to standards. *NASSP Bulletin, 84*(620), 37-48.

Kendall, J. S., & Marzano, R. J. (2000). *Content knowledge: A compendium of standards and benchmarks for K–12 education* (3rd ed.). Reston, VA: Association for Supervision and Curriculum Development.

Kendall, J. S., Schoch-Roberts, L., & Young-Reynolds, S. (2000). *A distillation of subject-matter content for the subject areas of geography and history* (Tech. Rep.). Aurora, CO: Mid-continent Regional Educational Laboratory.

Kendall, J. S., Snyder, C., Schintgen, M., Wahlquist, A., & Marzano, R. J. (1999). *A distillation of subject-matter content for the subject areas of language arts, mathematics, and science* (Tech. Rep.). Aurora, CO: Mid-continent Regional Educational Laboratory.

Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Laboratory Network Program.(1998). Taking stock of states' curriculum-based reform efforts: An interim report of the Laboratory Network Program's Curriculum, Learning, and Instruction Project. Aurora, CO: Mid-continent Regional Educational Laboratory.

Lerner, L. S. (1998, March). State science standards: An appraisal of science standards in 36 states. *Fordham Report, 2*(4).

Louisiana State Department of Education. (1997, May). *Social studies content standards.* Retrieved June 20, 2000 from the World Wide Web: http://www.lcet.doe.state.la.us/doe/asps/home.asp?I=CONTENT

Marzano, R. J., & Kendall, J. S. (1996). *A comprehensive guide to designing a standards-based district, school, or classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., & Kendall, J. S. (with Gaddy, B. B.). (1999). *Essential knowledge: The debate over what American students should know.* Aurora, CO: Mid-continent Regional Educational Laboratory.

Massachusetts Department of Education. (1997, February). *The English language arts curriculum framework.* Malden, MA: Author. Retrieved May 18, 1999 from the World Wide Web: http://www.doe.mass.edu/doedocs/frameworks/englishTOC.html

Missouri State Board of Education. (1996). *The show-me standards: Mathematics.* Retrieved May 18, 2001 from the World Wide Web: http://www.dese.state.mo.us/standards/math.html

Monroe, S. (2000). The state of state standards in geography. In C. E. Finn, Jr. & M. J. Pitrelli (Eds.), *The state of state standards 2000.* Washington, DC: The Thomas B. Fordham Foundation. Retrieved March 11, 2000 from the World Wide Web: http://www.edexcellence.net/library/soss2000/Standards2000.pdf

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

The National Education Standards and Improvement Council. (1993). Promises to keep: Creating high standards for American students. Report on the review of educational standards from the Goals 3 and 4 Technical Planning Group to the National Education Goals Panel. Washington, DC: National Goals Panel.

New York State Education Department. (1996, March). *Learning standards for mathematics, science, and technology* (Rev. ed.). Albany: Author.

North Dakota Department of Public Instruction. (Spring 1999). *North Dakota standards and benchmarks: Content standards: Mathematics*. Retrieved May 20, 2001 from the World Wide Web: http://www.dpi.state.nd.us/standard/content/math.pdf

Ohio Department of Education, Division of Elementary and Secondary Education. (1990, November). *Model competency-based mathematics program*. Columbus, OH: Author.

Olson, Lynn. (1998, April 15). An "A" or a "D": State rankings differ widely. *Education Week*. Retrieved May 10, 2001 from the World Wide Web: http://www.edweek.org/ew/vol-17/31stand.h17

Pennsylvania State Board of Education. (1999, January). *Academic standards for mathematics*. Retrieved May 20, 2001 from the World Wide Web:http://www.edexcellence.net/standards/

Raimi, R.& Braden, L. (1998). State mathematics standards: An appraisal of math standards in 46 states, the District of Columbia, and Japan. *Fordham Report* 2(3). Retrieved October 20, 1999 from the World Wide Web: http://www.edexcellence.net/standards/math/math.htm

Rhode Island Department of Education. (1996, August 14). *Rhode Island science framework*. Providence, RI: Author. Retrieved June 17, 1999 from the World Wide Web: http://instruct.ride.ri.net/doehome/scope.html

Saxe, D. W. (2000). The state of state standards in history. In C. E. Finn, Jr. & M. J. Pitrelli (Eds.), *The state of state standards 2000*. Washington, DC: The Thomas B. Fordham Foundation. Retrieved March 11, 2000 from the World Wide Web: http://www.edexcellence.net/library/soss2000/Standards2000.pdf

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U.S. science and mathematics. Executive summary*. Lansing, MI: U.S. National Research Center for the Third International Mathematics and Science Study, Michigan State University. Retrieved December 1, 1999 from the World Wide Web: http://ustimss.msu.edu/splintrd.pdf

Singapore (1994). *Primary Mathematics Guide*. Singapore: Federal Publications.

South Carolina State Department of Education. (2000, March). *Curriculum standards: Social studies*. Retrieved June 20, 2000 from the World Wide Web: http://www.state.sc.us/sde/educator/ standard/socstd/index.html

South Dakota State Board of Education. (1998, December). *South Dakota mathematical standards*. Retrieved May 20, 2001 from the World Wide Web: http://www.state.sd.us/deca/ContentStandards/math/math.htm

Utah State Office of Education. (1994). *Core curriculum standards: Mathematics*. Salt Lake City, UT: Author. Retrieved June, 18, 1999 from the World Wide Web: http://www.uen.org/ cgi-bin/ websql/lessons/query_lp.hts?corearea=2&area=1

West Virginia Department of Education. (1995, June). *West Virginia programs of study: Instructional goals and objectives*. Charleston, WV: Author.

Wisconsin Department of Public Instruction. (1999). *Wisconsin's model academic standards for English language arts*. Madison, WI: Author. Retrieved May 18, 1999 from the World Wide Web: http://www.dpi.state.wi.us/standards/elaintro.html

Wyoming Department of Education. (1998, June). *Wyoming mathematics content and performance standards*. Retrieved May 20, 2001 from the World Wide Web: http://www.k12.wy.us/publications/math.pdf